

Infliacijos prognozavimas faktoriniais modeliais

Dovilė ŠIURKUTĖ, Audronė JAKAITIENĖ

Vilniaus Gedimino Technikos universitetas

Saulėtekio al. 11, LT-10223 Vilnius

el. paštas: siurkute@yahoo.com; audrone.jakaitiene@ktl.mii.lt

Santrauka. Šiame darbe pritaikyti didelės apimties faktoriniai modeliai Lietuvos vartotojų ir gamintojų mėnesinei inflacijai prognozuoti. Ištirtas sudarytų faktorių modelių pranašumas, lyginti su atsitiktinio klajojimo ir pirmos eilės autoregresijos modeliais. Tyrime naudojamos 147 mėnesinio dažnumo laiko eilutės nuo 1996 m. iki 2007 m. Gauta, kad gamintojų inflaciją Lietuvoje tiksliausiai prognozuoja faktoriiniai modeliai. Vartotojų inflaciją – atsitiktinio klajojimo modelis, tačiau faktorinis modelis yra pranašesnis, palyginti su pirmos eilės autoregresijos modeliu, atliekant trumpalaikę prognozę. Ištyrus faktorinius modelius, sudarytus pagal duomenų grupes, gauta, kad vartotojų inflaciją Lietuvoje geriausiai atspindi kainų duomenys, o modelio maksimaliam prognozės tikslumui gauti pakanka panaudoti vieną faktorių.

Raktiniai žodžiai: faktorinė analizė, inflacija, pagrindinių komponentų metodas.

Ivadas

Inflacija yra viena opiausių šiuolaikinės makroekonomikos problemų, todėl kainų augimui analizuoti, modeliuoti ir prognozuoti pastaruoju metu visame pasaulyje skiriama nemažai tyrimų, taikomi įvairūs ekonometrinės analizės metodai, ne išimtis ir Lietuva. Inflacijos prognozei taikyti tiesinės regresijos modeliai ir įprastiniai laiko eilučių modeliai [11], o kiek vėliau ir pažangūs laiko eilučių modeliai: vektorinės autoregresijos modelis, daugiamatis vektorinis paklaidų korekcijos modelis [5]. Tačiau nuolatinis duomenų kiekio augimas leidžia pereiti ir prie modelių, skirtų dideliame duomenų kiekiui analizuoti, t. y. faktorinės analizės. Vakarų šalyse faktoriniai modeliai inflacijai prognozuoti buvo pritaikyti gana plačiai ([1,4,6,9,10] ir kt.). Rytų Europoje faktorinės analizės panaudojimas makroekonominiais rodikliais prognozuoti pasiteisino ([2,3]), todėl tikėtina, kad ir Lietuvoje ši analizė bus rezultatyvi.

Darbe pritaikyti didelės apimties faktoriniai modeliai Lietuvos vartotojų ir gamintojų mėnesinei inflacijai prognozuoti. Faktoriniais modeliais apskaičiuotos prognozės lyginamos su atsitiktinio klajojimo ir pirmos eilės autoregresijos modelių prognozėmis.

1. Matematinis faktorinės analizės modelis

Tarkime, kad X_i – stebimo kintamojo reikšmių vektorius, F_j – bendrųjų latentinių faktorių vektorius, čia $i = 1, \dots, p$ ir $j = 1, \dots, k$, tada

$$X_i = \sum_{j=1}^k \lambda_{ij} F_j + e_i, \quad (1)$$

čia daugikliai λ_{ij} – faktorių svoriai, kuriuos galima apibrėžti kaip kovariaciją tarp X_i ir F_j , o dėmuo e_i – specifinis latentinis faktorius. Pritaikius faktorinės analizės metodus, yra randami faktorių svoriai, bendrumai ir specifiškumai. Darbe faktoriams rasti naudojamas pagrindinių komponentių metodas. Pagrindinių komponentių metodu randami (1) parametrai λ ir F , tokie, kad maksimuotų pradinių kintamųjų X paaiškintos variacijos dalį, juos pakeitus k faktoriais.

2. Tyrimo duomenys ir jų tinkamumas faktorinei analizei

Darbe naudojamos 147 mėnesinio dažnumo laiko eilutės, kurios apima laikotarpį nuo 1996 m. sausio mėn. iki 2007 m. gruodžio mėn. Faktorinio modelio įvertinimui naudojami duomenys iki 2004 m. gruodžio mėn., o infliacijos prognozė vertinama laikotarpyje nuo 2005 m. sausio mėn. iki 2007 m. gruodžio mėn. Norint nustatyti skirtingų duomenų įtaką infliacijai Lietuvoje prognozuoti, duomenys yra suskirstomi į keturias duomenų grupes: realūs duomenys (71), finansiniai duomenys (22), kainų duomenys (31) ir pasitikėjimo rodikliai (23). Šiam tyrimui buvo naudotos tik viešai skelbiamos Statistikos departamento prie Lietuvos Respublikos vyriausybės, Lietuvos banko ir Europos komisijos laiko eilutės¹. Lietuvos vartotojų mėnesinė infliacija apskaičiuota naudojant suderintą vartotojų kainų indeksą, o atitinkamai gamintojų infliacijai naudotas gamintojų kainų indeksas.

Prieš taikant faktorinę analizę, tikrinamas duomenų tinkamumas, taikant Kaizerio–Mejerio–Olkinio (KMO) matą:

$$KMO = \frac{\sum \sum_{i \neq j} r_{ij}^2}{\sum \sum_{i \neq j} r_{ij}^2 + \sum \sum_{i \neq j} \tilde{r}_{ij}^2}, \quad (2)$$

čia r_{ij} – kintamųjų X_i ir X_j koreliacijos koeficientas; \tilde{r}_{ij} – kintamųjų X_i ir X_j dalinės koreliacijos koeficientas. Naudojant visus turimus duomenis, $KMO_{147} = 0,027$ ir kadangi $KMO_{147} < 0,6$, tai faktorinė analizė nerezultatyvi, t. y. kintamųjų porų koreliacija nėra paaiškinama kitais kintamaisiais. Todėl iš pradinių kintamųjų sąrašo pašalinsime nepriklausomus kintamuosius. Kiekvieno kintamojo tinkamumo matas apskaičiuojamas pagal formulę:

$$MSA_i = \frac{\sum_{j \neq i} r_{ij}^2}{\sum_{j \neq i} r_{ij}^2 + \sum_{j \neq i} \tilde{r}_{ij}^2}. \quad (3)$$

Šalinamas tas kintamasis, kurio tinkamumo matas MSA yra mažiausias ir KMO matas perskaičiuojamas. Procedūra kartojama tol, kol KMO mato reikšmė viršys 0,6. Atlikus

¹Dėl didelio duomenų kiekio, laiko eilučių sąrašas nepriedamas, bet jį galima gauti tiesiogiai susisiekus su autorėmis.

duomenų tinkamumo faktorinei analizei testą, paaiškėjo, kad laiko eilučių tinkamų faktorinei analizei yra 90. Toliau tik šios laiko eilutės naudojamos faktorinio modelio sudarymui.

Prieš atliekant faktorinę analizę, duomenys buvo pakoreguoti pagal sezoniskumą, naudojant Tramo/Seats [7], tada laiko eilutėms pritaikytas pirmos eilės diferencijavimas ir standartizavimas, atimant laiko eilutės vidurkį ir padalinant iš dispersijos.

3. Prognozavimas taikant faktorinį modelį

Darbe atliktas prognozavimas remiasi Stocko ir Watsono [8] pasiūlyta prognozavimo, naudojant faktorinius modelius, schema. Infliacijos prognozavimas susideda iš dviejų etapų: pirmiausia pagrindinių komponentų metodu yra išskiriami pagrindiniai faktoriai iš stebimų duomenų rinkinio X , tada šie faktoriai yra naudojami pasirinkto kintamojo y prognozavimui. Taikomas h -žingsnių į priekį prognozavimas, kai h – prognozavimo horizontas. Prognozavimo uždavinys apibrėžiamas taip:

$$y_{t+h} = \alpha_h + \beta_h \hat{F}_t + \epsilon_{t+h}. \quad (4)$$

Lygties koeficientai $\hat{\alpha}_h$ ir $\hat{\beta}_h$ įvertinami naudojant mažiausių kvadratų metodą (MKM). Prognozė $y_{T+h|T}$ yra apibrėžiama taip:

$$y_{T+h|T} = \hat{\alpha}_h + \hat{\beta}_h \hat{F}_T. \quad (5)$$

Norint pagrįsti modelių pranašumą, darbe sudarytų faktorinių modelių prognozės rezultatai yra lyginami su dviejų modelių – atsitiktinio klajojimo ir pirmos eilės autoregresijos – prognozių tikslumu. Atsitiktinio klajojimo modelis pasirenkamas, nes tai vienas paprasčiausių ir nereikalaujantis didelių sąnaudų prognozavimo modelis. Infliacijai prognozuoti dažnai naudojami autoregresijos modeliai, kurie yra gana tikslūs.

Tarkime, kad Y_t – stacionarus kintamasis, tuomet atsitiktinio klajojimo procesas apibrėžiamas taip:

$$y_t = y_{t-1} + \epsilon_t, \quad (6)$$

čia Y_t – kintamojo Y stebiniai laiko momentu $t = 1, \dots, T$, o ϵ_t – baltasis triukšmas.

Tada stacionariam kintamajam Y_t pirmos eilės autoregresinis procesas (žymimas AR(1)) apibrėžiamas taip:

$$y_t = \alpha + \phi y_{t-1} + \epsilon_t, \quad (7)$$

kur α yra konstanta, ϕ – autoregresijos proceso parametras.

Prognozė, gauta naudojant faktorinius modelius, lyginama su prognozėmis, gautomis naudojant atsitiktinio klajojimo modelį (toliau – RW) ir pirmos eilės autoregresijos modelį. Kad būtų galima palyginti gautas prognozes, atliekamas prognozavimas už imties ribų (angl. out of sample forecast). Modelių palyginimui naudojamas prognozavimo paklaidos matas RMSE (angl. Root Mean Squared Error):

$$RMSE = \sqrt{\frac{1}{T} \sum_{t=1}^T (\hat{y}_t - y_t)^2}, \quad (8)$$

kur T – stebinių skaičius, \hat{y}_t – prognozuojama reikšmė, o y_t – stebėta reikšmė, kai $t = 1, \dots, T$. Apskaičiavus kiekvieno modelio paklaidos matą RMSE, skaičiuojami faktorinio ir autoregresijos modelių santykiniai paklaidos matai

$$RRMSE_F = \frac{RMSE_F}{RMSE_{RW}}, \quad RRMSE_{AR(1)} = \frac{RMSE_{AR(1)}}{RMSE_{RW}},$$

kur $RRMSE_F$ – faktorinio modelio santykinis paklaidos matas; $RRMSE_{AR(1)}$ – autoregresijos modelio santykinis paklaidos matas; $RMSE_F$ – faktorinio modelio paklaidos matas; $RMSE_{RW}$ – atsitiktinio klajojimo modelio paklaidos matas; $RMSE_{AR(1)}$ – autoregresijos modelio paklaidos matas. Apskaičiavus ir palyginus sudarytų modelių santykinius paklaidos matus, galima daryti išvadas apie modelių tinkamumą infliacijos prognozavimui, lyginant su atsitiktinio klajojimo metodu, t.y. jeigu $RRMSE$ yra mažesnis už vienetą, tai atitinkamo modelio prognozė yra tikslesnė nei naivi atsitiktinio klajojimo prognozė.

1 ir 2 lentelėse pateikiami santykiniai paklaidos matavimų vartotojų ir gamintojų infliacijų 12 mėn. prognozei. Matome, kad mėnesinei vartotojų infliacijai prognozuoti tinkamiausias yra atsitiktinio klajojimo modelis, tačiau faktorinis modelis yra pranašesnis nei pirmos eilės autoregresinis modelis iki 8 mėn. Tiesa, atliekant prognozavimą su skirtingomis duomenų grupėmis, buvo gautas tikslesnis faktorinio modelio rezultatas vienam mėnesiui prognozuoti. Gamintojų kainoms prognozuoti faktoriniai modeliai yra tinkamiausi iš visų nagrinėtų, kai naudojami visų tipų duomenys.

4. Išvados ir pasiūlymai

Mažiausios paklaidos vartotojų infliacijai Lietuvoje prognozuoti gaunamos, naudojant atsitiktinio klajojimo modelį, tačiau faktoriniai modeliai yra pranašesni, palyginti su pirmos eilės autoregresijos modeliu, atliekant trumpalaikę prognozę (iki

1 lentelė. Vartotojų mėnesinės infliacijos trijų faktorių visų duomenų (F_{3-B}), vieno faktoriaus kainų duomenų (F_{1-K}) ir autoregresijos modelių santykiniai paklaidos matai

[illegible]

2 lentelė. Gamintojų mėnesinės infliacijos penkų faktorių F_5) ir autoregresijos modelių santykiniai paklaidos matai

[illegible]

8 mėnesių). Tiksliausios gamintojų infliacijos prognozės Lietuvoje gaunamos naudojant faktorinius modelius. Palyginus sudarytus faktorinius modelius pagal duomenų grupes, galima teigti, kad vartotojų mėnesinę infliaciją tiksliausiai prognozuoja faktorinis modelis, sudarytas tik iš kainų duomenų. Tuo tarpu gamintojų mėnesinės infliacijos prognozės tikslumui reikšmingos įtakos duomenų tipas neturi. Sudarius faktorinius modelius Lietuvos vartotojų ir gamintojų mėnesinei infliacijai prognozuoti iš visų duomenų, paaiškėjo, kad mažiausios vartotojų infliacijos prognozės paklaidos gaunamos naudojant modelį, sudarytą iš trijų faktorių, o mažiausios gamintojų infliacijos paklaidos gaunamos naudojant modelį, sudarytą iš penkių faktorių. Pastebėsime, kad darbe buvo prognozuojama bendra vartotojų kainų infliacija, kuri apima ir administracijos sprendimus, kaip kad reguliuojamos kainos ar akcizų kėlimas, kuriems prognozavimo modelis yra nereikalingas. Todėl darbą galima būtų testuoti pritaikant faktorinius modelius vartotojų infliacijai be administruojamų kainų prognozuoti.

Literatūra

1. M.J. Artis, A. Banerjee, M. Marcellino. Factor forecasts for the UK. *Journal of Forecasting*, 24(4):279–298, 2005.
2. A. Banerjee, M. Marcellino, I. Masten. Forecasting macroeconomic variables for the new member states of the European Union. *ECB Working Paper Series*, 482, 2005.
3. K. Barhoumi, S. Benk, R. Cristadoro and others. Short-term forecasting of GDP using large monthly datasets: a pseudo real-time forecast evaluation exercise. *ECB Occasional Paper Series*, 84, 2008.
4. C. Cheung, F. Demers. Evaluating forecasts from factor models for Canadian GDP growth and core inflation. *Bank of Canada Working Paper*, 8, 2007.
5. A. Čuvak, Ž. Kalinauskas. Daugiamačių regresijos modelių naudojimas infliacijai modeliuoti. *Liet. matem. rink.*, 47(spec. nr.):434–441, 2007.
6. S. Eickmeier, C. Ziegler. How successful are dynamic factor models at forecasting output and inflation? A meta-analytic approach. *Journal of Forecasting*, 27(3):237–265, 2008.
7. V. Gomez, A. Maravall. Programs TRAMO and SEATS. Instruction for User (Beta Version: september 1996). *Banco de Espana Working Paper*, 9628, 1996.
8. J.H. Stock, M.W. Watson. A comparison of linear and nonlinear univariate models for forecasting macroeconomic time series. *NBER Working Paper*, 6607, 1998.
9. J.H. Stock, M.W. Watson. Forecasting inflation. *Journal of Monetary Economics*, 44(2):293–335, 1999.
10. J.H. Stock, M.W. Watson. Macroeconomic forecasting using diffusion indexes. *Journal of Business and Economic Statistics*, 20(2):147–162, 2002.
11. I. Vetlov. Lietuvos infliacijos inercijos analizė. *Pinigų studijos*, 3:5–16, 2000.

SUMMARY

D. Šiurkutė, A. Jakaitienė. Forecasting inflation in short-term using factor models

In this paper we apply the factor models to produce short-term forecasts for Lithuanian consumer and producer inflation. The factor models are compared with a random walk and the first order autoregression models. In this research work are used 147 time series publicly available at monthly frequency from 1996 until 2007. Research shows that, according Kaiser-Meyer-Olkin test, the most of these series (61 percent) are suitable for the factor analysis. The best forecast of producer price inflation is obtained by employing the factor model. The best forecast of consumer price inflation (according harmonised index of consumer prices) is made by employing the random walk model, but the factor model is better as the first order autoregression model at short-term forecast horizon. Price data best describes consumer inflation for Lithuania and a single factor is needed to produce maximum accuracy of the model.

Keywords: factor analysis, inflation, principal component method.